

# IS PSYCHOSIS CAUSED BY DEFECTIVE DISSOCIATION? AN ARTIFICIAL LIFE MODEL FOR SCHIZOPHRENIA

Alessandro Fontana

## Abstract

Both neurobiological and environmental factors are known to play a role in the origin of schizophrenia, but no model has been proposed that accounts for both. This work presents a functional model of schizophrenia that merges psychodynamic elements with ingredients borrowed from the theory of psychological traumas, and evidences the interplay of traumatic experiences and defective mental functions in the pathogenesis of the disorder. Our model foresees that dissociation is a standard tool used by the mind to protect itself from emotional pain. In case of repeated traumas, the mind learns to adopt selective forms of dissociation to avoid pain without losing touch with external reality. We conjecture that this process is defective in schizophrenia, where dissociation is either too weak, giving rise to positive symptoms, or too strong, causing negative symptoms.

## 1 Introduction

Mental disorders can be described at multiple levels: psychological, neurobiological, genetic. In this work we deal with another level, that we call functional, situated between the psychological and the neurobiological levels. The functional level aims to characterize the structure of the system under investigation and the function of each component, regardless of its neurobiological implementation. This is the scientific approach of Artificial Life: study life as it could be, to understand life as it is. This work presents a functional model of schizophrenia, which is used to interpret the psychological-symptomatic aspects of the disorder and to account for the neurobiological substrate.

Both genetic and environmental factors, such as traumatic experiences, are known to play a role in the origin of schizophrenia, but few models have been proposed that account for both. The model we are going to propose merges psychodynamic elements with ingredients borrowed from the theory of psychological traumas, and evidences the interplay of traumatic experiences and defective mental functions in the pathogenesis of schizophrenia. A thorough characterization of this complex condition is outside the scope of this work: our model, based on a simplification of reality, has the purpose to interpret some limited aspects of mental functioning.

Among theories of schizophrenia, much attention has been dedicated to the “dopamine hypothesis” (Howes and Kapur, 2009). Psychosis is hypothesized to be caused by the phenomenon of “aberrant salience” (Kapur, 2003), determined by an excess of neurotransmitter dopamine in the limbic system. This theory is supported by the fact that amphetamines, which trigger the

release of dopamine, may induce hallucinations and delusions (Laruell et al., 1996) and that antipsychotic drugs block the D2 dopamine receptor. However, large genetic studies were unable to detect genes linked to dopamine transmission (Alexis et al., 2016).

Another theory links schizophrenia to a malfunction of glutamate NMDA receptors which, unlike dopamine receptors, are distributed across the whole brain (Javitt, 2010). This theory is based on the observation that some drugs (such as PCP) which block the NMDA receptor, produce symptoms similar to those of schizophrenia. Unlike amphetamines, such drugs would be able to reproduce the full spectrum of schizophrenia symptoms (positive, cognitive, negative).

Connections between neurons are created at a very high pace during postnatal development, to reach their maximum density around age six. Many connections are subsequently eliminated during the “pruning” phase which, for prefrontal cortex, peaks in late adolescence, a period that corresponds to schizophrenia onset. A recent genetic study (Sekar et al., 2016) found a correlation with a gene coding for a receptor that targets synapses for destruction by the immune system. According to the study’s authors, the gene variant would cause excessive pruning and lead to schizophrenia. However, the defective gene is only present in a small minority of patients.

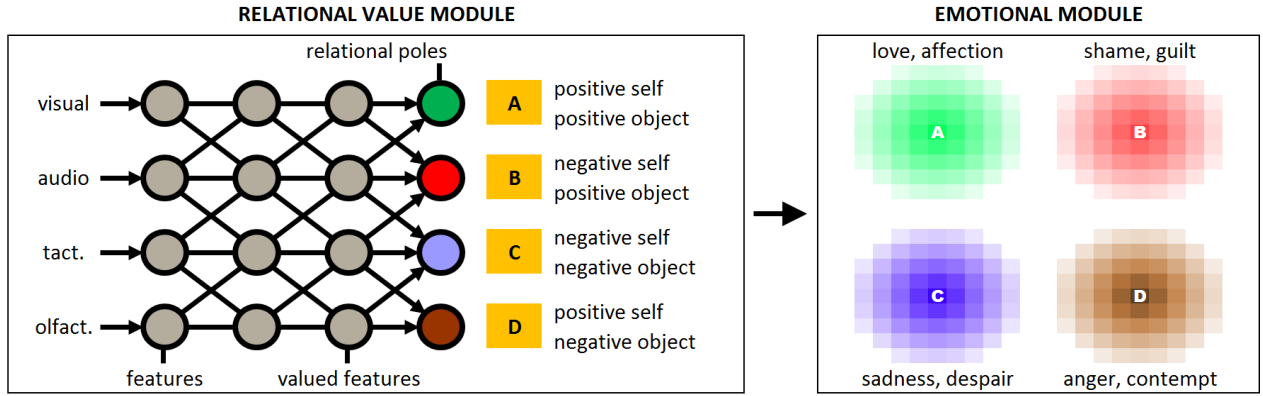
A history of childhood traumas is reported to be a risk factor for schizophrenia. The content of positive symptoms is coherent with low self-esteem and negative affect, which are the common outcome of adverse childhood experiences (Schaefer and Fisher, 2011). The effect of traumas could be mediated and amplified by enhanced susceptibility (Horan and Blanchard, 2003; Cohen and Docherty, 2004). The rest of the paper is organized as follows: the model is introduced in section 2 and 3, and used in section 4 to obtain an interpretation of some aspects of schizophrenia; section 5 draws the conclusions and outlines future research directions.

## 2 A model for the mind

### Value of features

External reality can be conceived a set of **situations**, each characterized through a list of active **features**. Examples of simple perceptual features are: shape, orientation, color. Examples of more abstract features are: “black ducks”, “soccer players”, “to be a teacher”. Feature activation and deactivation is a continuous process, driven by perceptual features fed from sensory stimuli and propagated to more abstract ones in real time. This occurs on a fast time scale, as the mind “navigates” through everyday life. For instance, if a person is walking on the beach, the features encoding the concepts of “sand” and “sun” may be active, while features such as “office chair” and “computer screen” are likely to be inactive.

We postulate that features are characterized by a property called **value**, that can be either positive or negative. Features such as “honesty” and “health”, for example, are considered positive, while features such as “deception” and “illness” are generally perceived as negative. The determination and change of feature value happens by association: if a new feature of unknown value is presented in association with positive (negative) features, it will assume a positive (negative) connotation. Since a feature in general takes part in many associations, its value will be



**Figure 1:** Relational judgements and emotions. The relational module (on the left) adds value to features and selects one of four relational poles. Each pole corresponds to a different combination of states for the self and for the object, and is associated to specific emotions, produced by the emotional module (on the right).

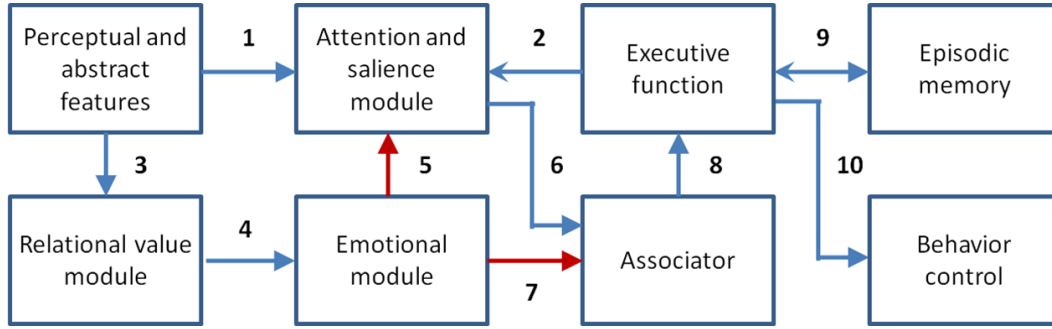
determined by the combined effect of all associations. Feature value is a long term property, expected to change on a slower time scale.

### Relational value module and emotions

Social and hierarchical judgements play a central role in human behavior. Emotions are often produced as a reaction to the actions of another person within a relational context, and depend on the appraisal of who is superior and who is inferior, who is right and who is wrong. Our assumption is that a specific **relational value module** is dedicated to managing relational judgements (Fig. 1). This module processes active features fed from the perceptual apparatus and assigns a value to two key symbols: the self and the object. As a result, it selects one of four possible **relational poles**, each associated to specific emotions, produce by an **emotional module**:

- A : positive self and positive object  
(the self feels love and affection towards a rewarding object)
- B : negative self and positive object  
(the self feels pain and shame in front of a superior /humiliating object)
- C : negative self and negative object  
(the self feels sadness or despair when self and object are both worthless)
- D : positive self and negative object  
(the self feels anger towards a guilty object or despises an inferior object)

Therefore, our assumption is that relational emotions are mediated by a relational-hierarchical judgement. The association of shame, guilt and sadness to a lower social status is confirmed by numerous studies (Stevens and Price, 1996). Also the association of anger to a higher social status has been recognized and studied (Tiedens, 2011). There are other emotions, such as fear, that appear to be more primordial and independent of social appraisal.



**Figure 2:** Functional modules of the mind.

## A model for the mind

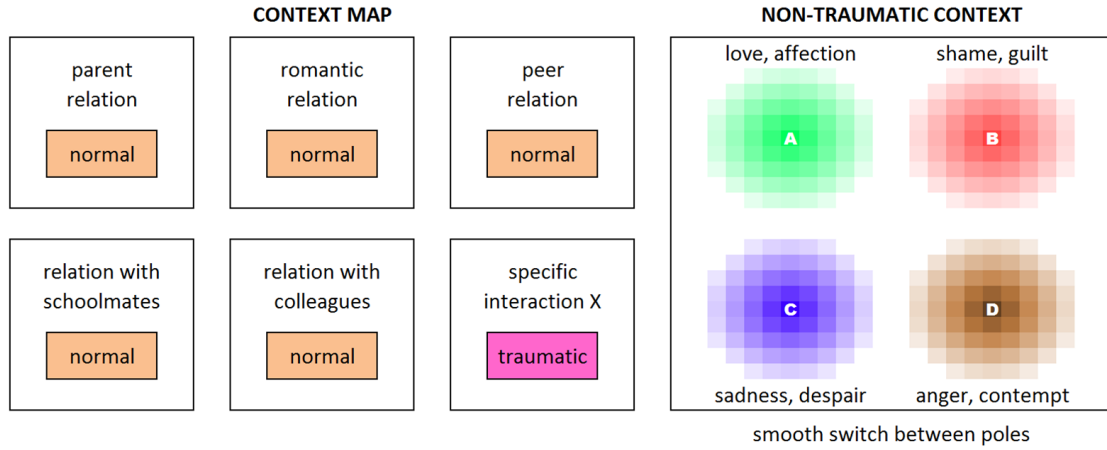
The mind can be represented as a set of interacting functional modules (Fig. 2). **Perceptual and abstract features** are created from the input flow and fed (link 1 in the figure) to the **attention and salience module**, which “adds” salience. Adding salience means simply paying more attention to features relevant for the task at hand. The information about the current task is provided (2) from the **executive module**, which has objectives and a plan to reach them.

Perceptual and abstract features are also fed in parallel (3) to the **relational value module** which, as described in the previous paragraph, adds value to active features and computes the value of the self and of the object. This enriched information is passed (4) to the **emotional module**, that responds with an appropriate emotion.

Emotions are fed (5) to the salience module, where they contribute to define the salience of active features. Features which contribute most to the elicited emotion are given more salience. If, for instance, the person feels ashamed because he/she has bat ears, the feature “ears” receives high salience. Therefore, our model foresees that salience is given to features based on goals and based on emotions.

Salient features and emotions are fed (6,7) to the associator, which groups them together and passes the link (8) to this record to the executive function. This module compares the current association with the content of **episodic memory** (9) and to its goals. Based on this, it stores the current record in memory and directs behavior through the **behavior control** module (10).

Based on current neurophysiological evidence, the processing of perceptual and abstract features takes place in the cortex (e.g., occipital cortex for visual features). The relational and emotional modules could be localized in the medial prefrontal, cingulate and insular cortex, regions critical for social interaction (Palmiter, 2008; Etkin et al., 2011; Gu et al., 2013). The executive function may also be implemented in the prefrontal cortex, while the salience module may correspond to the striatum, a region involved in reward mechanisms (Martin-Soelch et al., 2001). The associator may correspond to the hippocampus, a structure essential for memory formation. Motor control is implemented in the motor cortex, while episodic memory is distributed across the whole cortex.



**Figure 3:** Normal condition. In the non-pathological case traumatic contexts are very few and limited in scope (left panel). The mind can switch smoothly between relational poles in each context, thanks to an intact relational module. No pole requires dissociation: the repertoire of emotions is fully accessible (right panel).

### 3 Normal and traumatic contexts

#### Normal functioning

From a relational perspective, life situations can be grouped in different “contexts”, such as “parent relation”, “peer relation”, “relation with schoolmates” (Fig. 3-left). Each context contains situations that are similar and require a similar approach. Contexts can be represented on a plane, in which points correspond to individual situations (Fig. 3-right), characterized by different combinations of active features.

Each point belongs to the “zone of influence” of one of the four relational poles. Each pole originates from a point representing the most prototypical situation associated to the pole, and extends towards less prototypical situations. The epicenter of Pole D, for instance, may correspond to a situation-point characterized by an object behaving very dishonestly, eliciting a very strong anger, while points further away may be characterized by a better behavior of the object. In normal conditions, almost all contexts are “normal” (Fig. 3-left) and the mind can switch smoothly between poles, during its navigation through everyday life.

Emotional processing is complemented by salience generation. Features are passed to the salience module, which attaches importance to them based on their degree of contribution to the emotion. Salience is given to features also based on their relevance for the task currently pursued by the executive function. In other words, the salience module is telling the mind: “you are not far from emotion E, which mostly depends on feature X1” (emotional salience) or “to reach objective O on your agenda, you should pay attention to feature X2” (motivational salience). The idea that neurons in the limbic system encode two kinds of salience (emotional and motivational) is not new (Bromberg-Martin et al., 2010).

## Traumatic functioning

In a traumatic condition, one or more important contexts are traumatic (Fig. 4-left). The relational poles that can be involved in a trauma are those indicated with B and C in Fig. 4-right, in which the self is negative (**negative poles** for brevity). Pole B consists of a negative self and a positive object. This may correspond to a situation of physical or verbal abuse. Pole C consists of a self and an object which are both negative and worthless. This may correspond to a situation in which all family members are exposed to natural disasters, or poverty.

In case of trauma, the emotional values elicited are too high and the functioning of the relational module becomes critical. We assume that the level of emotion  $E$  is proportional to the difference between the object value  $V_{obj}$  and the self value  $V_{self}$  in case B, and to the difference between a reference value  $V_{ref}$  and the self value in case C. Self and object values are in turn determined by the sum of the values of all active features associated to them,  $V_{s_i}$  and  $V_{o_i}$ . In formulas ( $S$  is the sigmoid function, which prevents emotional values from exceeding a maximum threshold):

$$\text{Pole B: } E = S(V_{obj} - V_{self})$$

$$\text{Pole C: } E = S(V_{ref} - V_{self})$$

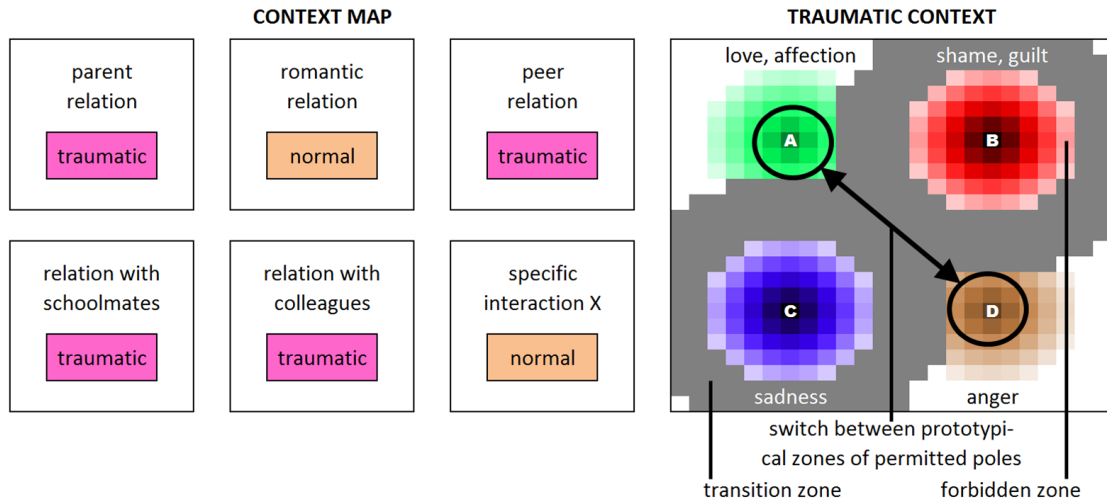
$$V_{self} = S(\sum_i V_{s_i}); V_{obj} = S(\sum_i V_{o_i})$$

When the value of the active features associated to the self is very low, the self is dragged to a very low value and the trauma occurs. If a person with bat ears, big nose and thin lips thinks that these features are very negative and he/she gets criticized or made fun of for them, the trauma takes place (of course this is just an example: real traumas are often more complicated and need to be repeated to cause damage).

Upon the occurrence of a trauma, the associator stops working: this corresponds to the phenomenon of **dissociation**, defined as the distortion, limitation or loss of the normal associative links between perceptions, emotions, thoughts and behavior. Dissociation can take the form of mental “black-out”, depersonalization (feeling of separation from one’s body), derealization (feeling of being detached from the world), selective amnesia and emotional detachment (Lanius, 2015; Radovic and Radovic, 2002).

In case of trauma, the adoption of dissociation makes it possible for the mind to stay on negative poles, excluding the awareness of intolerable thoughts and emotions from consciousness. However, the disconnection of aspects of perceived reality may hide potential dangers and have a high cost. Therefore, the mind tries to avoid the negative poles B and C, and heads towards the positive poles A and D, where perception of reality is not restricted. The space of a traumatic context can be divided into three zones (Fig. 4-right): the “forbidden zone”, an area near a traumatic pole, which cannot be accessed in a non-dissociated state; the “free zone”, an area far from all traumatic poles (or near the positive poles A and D); the “transition zone”, a safety belt around the forbidden zone.

The mind stays in the free zone around pole A as long as conditions are prototypical for pole A, i.e. as long as the relation with the object is perfect, full of trust, mutual respect, etc. As conditions depart from their prototypical values for pole A and drift into the transition zone,



**Figure 4:** Traumatic condition. In this case several contexts are traumatic (left). In a traumatic context (right), pole B and/or pole C are characterized by too high emotional levels, and are inhibited. When the mind happens to be in one of these poles (forbidden zone), dissociation kicks in. To avoid dissociation, the mind oscillates between poles A and D, staying in each pole as long as the situation is prototypical for the pole. When the situation drifts out of the free zone into the transition zone (shown in grey), splitting symptoms appear, and the mind switches to the free zone of the other permitted pole.

the mind switches abruptly to the free zone around pole D. When conditions are not perceived as prototypical for pole D, the mind jumps back to pole A, and the cycle repeats itself. This corresponds to the defense mechanism of **splitting**, defined as the inability to integrate positive and negative aspects of self and others. The result is a characteristic view of the world in “black and white” (Perry et al., 2013).

The transition zone is a safety belt built around the forbidden zone, characterized by high emotional level. We hypothesize that this increases the emotional salience of the features involved in the trauma (link 5 in Fig. 2), which may represent the person’s defects criticized. As a result, these features become the focus of attention and appear distorted and magnified. The distortion of defects serves the purpose to alert the executive function that the current situation is close to a traumatic point, and gives an indication of the features that need to be closely monitored. If the feature criticized is “bat ears”, when looking at the mirror the person will see his/her ears magnified and more protruded, like in a caricature. Sometimes, despite all efforts, the mind cannot avoid to fall back into a traumatic pole: when this happens, it resorts to dissociation to protect itself from pain.

## 4 Interpretation of schizophrenia

As already stated, both genetic and environmental factors play a role in schizophrenia. On the environmental side, childhood traumas involving caregivers are often present in the history of persons with schizophrenia (Schaefer and Fisher, 2011). As a result, the functioning of the relational module may be impaired in key contexts such as “parent relation” and “romantic relation”

(Fig. 4). In such contexts poles B and C are difficult to access, and the mind is forced to oscillate between poles A and D. This corresponds to splitting, which is the core defense mechanism employed by persons with borderline personality disorder, as well in a variety of other disorders (Perry et al., 2013).

We have seen how, in the transition zone of a traumatic context, features involved in the trauma are loaded with salience and as a result appear distorted and magnified, like in a caricature. This phenomenon is reminiscent of symptoms relevant to a wide range of mental disorders. Transitory perceptual distortions or “pseudo-hallucinations” are present in personality disorders (Gras et al., 2014). The difference between such distortions and schizophrenic hallucinations appears to be quantitative rather than qualitative. Feature distortion is also present in body dysmorphic disorder, which can either appear stand-alone (Phillips, 2004) or be responsible for the dysmorphic body image associated to eating disorders (among many other disorders) (Ruffolo et al., 2006).

We have seen that, when the mind falls in the forbidden zone (poles B or C) dissociation is used to avoid pain. We can hypothesize that, upon the first occurrence of a trauma, dissociation is total and involves both the cognitive and the emotional part (links 6 and 7 of Fig. 2 are both disconnected) This may correspond to the “freezing” behavior of children while interacting with their mothers, in case of “disorganized attachment” (Main and Solomon, 1986). This kind of dissociation provides an effective shield to emotional pain. On the other hand, a condition in which mental function is temporarily turned off exposes the individual to serious consequences in a potentially dangerous environment. The use of total dissociation as a defense mechanism is impractical, potentially fatal, and cannot be sustained for long periods of time.

Emotional dissociation is a more limited form of dissociation which, in a traumatic situation, excludes from consciousness the emotional component, while the rational part continues to be active. The subjective experience that accompanies this kind of dissociation may be represented by the phenomena of derealization and depersonalization, characterized by a diminished sense of reality. It is not unrealistic to assume that, in case of repeated traumas, the mind learns to replace total dissociation with emotional dissociation, as a compromise able to allow a higher level of functioning.

This is a common phenomenon. The first time that medical students see an operation, they may be shocked, but they rapidly get used to it. The same happens to doctors working in intensive care units, where death is a daily occurrence, or to the employees of a slaughterhouse. This happens to all of us, when we see poor people living in the street and pretend it is normal. We can think of this process as of a form of learning, in which the mind selects the smallest subset of reality that needs to be excluded from consciousness to avoid pain without losing touch with the “here and now”. This is obtained through a modulation of links 5 and 7 of Fig. 2 which can be attenuated in certain situations.

Our hypothesis is that the learning process that leads to emotional dissociation fails in schizophrenia: dissociation is either too weak or too strong. If it is too weak, links 5 and 7 are not attenuated, leaving the mind unprotected near the epicenter of a trauma. As a result, feature distortion caused by excessive salience proceeds unabated and becomes more pronounced: this is, in our opinion, the nature of hallucinations.



When the self and its associated features are strongly devalued, they are “responsible” for the guilt, shame or pain experienced (they are the major contributors to the emotion): as such, they are loaded with high salience. In such conditions, the disabling of the associator is necessary to avoid the establishment of unreliable associations, lacking the necessary statistical robustness. Otherwise salient features, such as the self, run the risk of being associated with random features which happen to co-occur. This explains the self-referential nature of delusions in schizophrenia.

Dissociation eventually intervenes, too strongly. When the dissociation wave hits the brain, it produces the full spectrum of negative symptoms: emotional detachment, anhedonia, lack of motivation and thought disorganization. The number and strength of links fed to the associator is reduced and the degree of reduction correlates with the severity of the symptoms observed.

As we said in section 2, the associator may correspond to the hippocampus, a brain region involved in memory formation. Interestingly, structural changes of the hippocampus are frequently observed in schizophrenia (Heckers and Konradi, 2010). The hippocampus of schizophrenic subjects is significantly smaller and seems to be less susceptible to the phenomenon of habituation (Williams et al., 2013), which could be necessary to develop selective and emotional dissociation. These observations led to the hypothesis of psychosis as a learning and memory problem (Tamminga, 2013), coherent with our theory.

The proposed theory requires the action of environmental factors (traumas), along with genetic factors (inability to learn selective dissociation) to produce the schizophrenic phenotype. While we can conjecture that the genetic predisposition is present at birth (although it could also develop later in life), the accumulation rate of traumatic events is expected to be constant in the course of life and to accelerate during adolescence, which is acknowledged to be a stressful period. This would explain why schizophrenia has its onset during adolescence.

## 5 Conclusions

The objective of this work was to merge the psychodynamic theory based on defense mechanisms and the theory of dissociation as a response to trauma, to produce a new interpretation of schizophrenia. The proposed theory highlights the interplay between environmental traumas and genetic predisposition for the production of the schizophrenic phenotype. Future work will be aimed to develop the model proposed and draw from it further insights for therapy.

## References

- Alexis, C. et al. (2016). Evaluating the dopamine hypothesis of schizophrenia in a large-scale genome-wide association study. *Schizophrenia Research*, 176:136–140.
- Bromberg-Martin, E., Matsumoto, M., and Hikosaka, O. (2010). Dopamine in motivational control: rewarding, aversive, and alerting. *Neuron*, 68(5):815–834.
- Cohen, A. and Docherty, N. (2004). Affective reactivity of speech and emotional experience in patients with schizophrenia. *Schizophrenia Research*, 69(1):7–14.

- Etkin, A., Egner, T., and Kalisch, R. (2011). Emotional processing in anterior cingulate and medial prefrontal cortex. *Trends Cogn Sci*, 15(2):85–93.
- Gras, A., Amad, A., Thomas, P., and Jardri, R. (2014). Hallucinations and borderline personality disorder: a review. *Encephale*, 40(6):431–438.
- Gu, X., Hof, P., Friston, K., and Fan, J. (2013). Anterior insular cortex and emotional awareness. *Journal of Comparative Neurology*, 521(15):3371–3388.
- Heckers, S. and Konradi, C. (2010). Hippocampal pathology in schizophrenia. *Current Topics in Behavioral Neurosciences*, 4:529–553.
- Horan, W. and Blanchard, J. (2003). Emotional responses to psychosocial stress in schizophrenia: the role of individual differences in affective traits and coping. *Schizophrenia Research*, 60(2-3):271–83.
- Howes, O. and Kapur, S. (2009). The dopamine hypothesis of schizophrenia: version iii: the final common pathway. *Schizophrenia Bulletin*, 35(3):549–562.
- Javitt, D. (2010). Glutamate as a therapeutic target in psychiatric disorders. *Nature Publishing Group*, 9(11):984–997.
- Kapur, S. (2003). Psychosis as a state of aberrant salience: a framework linking biology, phenomenology, and pharmacology in schizophrenia. *American Journal of Psychiatry*, 160:13–23.
- Lanius, R. (2015). Trauma-related dissociation and altered states of consciousness: a call for clinical, treatment, and neuroscience research. *European Journal of Psychotraumatology*, 6.
- Laruell, M. et al. (1996). Single photon emission computerized tomography imaging of amphetamine-induced dopamine release in drug-free schizophrenic subjects. *PNAS*, 93:9235–9240.
- Main, M. and Solomon, J. (1986). Discovery of an insecure disoriented attachment pattern: procedures, findings and implications for the classification of behavior. *Affective development in infancy*.
- Martin-Soelch, C. et al. (2001). Reward mechanisms in the brain and their role in dependence: evidence from neurophysiological and neuroimaging studies. *Brain Res Brain Res Rev*, 36(2-3):139–149.
- Palmiter, R. (2008). Dopamine signaling in the dorsal striatum is essential for motivated behaviors: lessons from dopamine-deficient mice. *Annals of the New York Academy of Sciences*, 1129:35–46.
- Perry, J., Presniak, M., and Olson, T. (2013). Defense mechanisms in schizotypal, borderline, antisocial, and narcissistic personality disorders. *Psychiatry Interpersonal and Biological Processes*, 76(1):32–52.

- Phillips, K. (2004). Body dysmorphic disorder: recognizing and treating imagined ugliness. *World Psychiatry*, 3(1):12–17.
- Radovic, F. and Radovic, S. (2002). Feelings of unreality: a conceptual and phenomenological analysis of the language of depersonalization. *Philosophy, Psychiatry and Psychology*, 9(3):271–279.
- Ruffolo, J. et al. (2006). Comorbidity of body dysmorphic disorder and eating disorders: severity of psychopathology and body image disturbance. *The International Journal of Eating Disorders*, 39(1):11–19.
- Schaefer, I. and Fisher, H. (2011). Childhood trauma and psychosis - what is the evidence? *Dialogues Clin Neuroscience*, 13(3):360–365.
- Sekar, A. et al. (2016). Schizophrenia risk from complex variation of complement component 4. *Nature*, 530:177–183.
- Stevens, A. and Price, J. (1996). *Evolutionary psychiatry: a new beginning*. New York: Routledge.
- Tamminga, C. (2013). Psychosis is emerging as a learning and memory disorder. *Neuropsychopharmacology Reviews*, 38:247.
- Tiedens, L. (2011). Anger and advancement versus sadness and subjugation: the effect of negative emotion expressions on social status conferral. *Journal of Personality and Social Psychology*, 80(1):86–94.
- Williams, L. et al. (2013). Reduced habituation in patients with schizophrenia. *Schizophr Res.*, 151(1-3):124–132.